2. Classic Cryptography Methods

2.1. Spartan scytale.

One of the oldest known examples is the Spartan scytale (scytale /skɪtəli/, rhymes with Italy, a baton). From indirect evidence, the scytale was first mentioned by the Greek poet Archilochus who lived in the 7th century B.C. (over 2500 years ago). The ancient Greeks, and the Spartans in particular, are said to have used this cipher to communicate during military campaigns.Sender and recipient each had a cylinder

(called a *scytale*) of exactly the same radius. The sender wound a narrow ribbon of parchment around his cylinder, then wrote on it *lengthwise*. After the ribbon is unwound, the writing could be read only by a person who had a cylinder of exactly the same circumference.
The following table illustrate the idea. Imagine that each column wraps around the dowel one time, that is that the bottom of one column is followed by the top of the next column.

Original message: *Kill king tomorrow midnight*

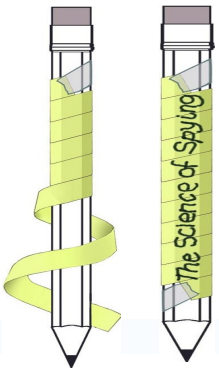
Wrapped message:

k	i	l	l	k	i	n	g
t	o	m	o	r	r	o	w
m	i	d	n	i	g	h	t

Encoded message: ktm ioi lmd lon kri irg noh gwt

The key parameter in using the scytale encryption is the number of letters that can be recorded on one wrap ribbon around the dowel. Above the maximum was 3, since there are 3 rows in the wrapped meassage. The last row was padded with blank spaces before the message was encoded. We'll call this the wrap parameter. If you don't know the wrap parameter you cannot decode a message.

Spartan skytale example with pensil and paper



2.2. Polybius square.

In cryptography, the Polybius square (205-123 B.C), also known as the Polybius checkerboard, is a device invented by the Ancient Greek historian and scholar Polybius..
The original square used the Greek alphabet, but can be used with any alphabet. In fact, it has also been used with Japanese hiragana (*see cryptography in Japan*). With the modern English alphabet, in typical form, it appears thus:

	1	**2**	**3**	**4**	**5**
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	O	R	S	T	U
5	V	W	X	Y	Z

Each letter is then represented by its coordinates in the grid. For example, "BAT" becomes "12 11 44". Because 26 characters do not quite fit in a square, we round down to the next lowest square number by combining two letters - I and J, usually. (Polybius had no such problem because the Greek alphabet he was using had 24 letters). Alternatively, we could add digits as well and get a 6 × 6 grid. Such a larger grid might also be used for the Cyrillic alphabet (of which the most common variant has 33 letters, though some have fewer, and some up to 37.)
Polybius did not originally conceive of his device as a cipher so much as an aid to telegraphy; he suggested the symbols could be signaled by holding up pairs of sets of torches. It has also been used, in the form of the "knock code", to signal messages between cells in prisons by tapping the numbers on pipes or walls. In this form it is said to have been used by nihilist prisoners of the Russian Czars, and also by American prisoners of war in the Vietnam War. Indeed it can be signaled in many simple ways (flashing lamps, blasts of sound, drums, smoke signals) and is much easier to learn than more sophisticated codes like the Morse code. However, it is also somewhat less efficient than the more complex codes.

2.3. Ceasar Cipher

Another early encrytption method is the Caesar Cipher. The Ceasar Cipher is an example of what is called a shift cipher. To encode a message, letters are replaced with a letter that is a fixed number of letters beyond the current letter. If the you run past the end of the alphabet you wrap around to the begining. To decode you shift backward, wrapping from the begining to the end if needed.
The essential piece of information for a Caesar Cipher is the shift parameter. Once this is known then a message may be encoded or decoded very easily.
The transformation can be represented by aligning two alphabets; the cipher alphabet is the plain alphabet rotated left or right by some number of positions. For instance, here is a Caesar cipher using a left rotation of three places (the shift parameter, here 3, is used as the key):

PT	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
CT	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

When encrypting, a person looks up each letter of the message in the plaintext (PT) line and writes down the corresponding letter in the ciphertext (CT) line. Deciphering is done in reverse.

Example: Plaintext: “Ankara” with 3 shift parameter will be “DQNAUD”.

Ciphertext: WKH TXLFN EURZQ IRA MXPSV RYHU WKH ODCB GRJ
Plaintext: the quick brown fox jumps over the lazy dog
The shift parameter is 11 in the example below.

The cipher is classed as a type of *monoalphabetic substitution*, as opposed to *polyalphabetic substitution*.
It is unknown how effective the Caesar cipher was at the time, but it is likely to have been reasonably secure, not least because most of Caesar's enemies would have been illiterate and others would have assumed that the messages were written in an unknown foreign language. There is no record at that time of any techniques for the solution of simple substitution ciphers. The earliest surviving records date to the 9th century works of Al-Kindi in the Arab world with the discovery of frequency analysis.

In the 19th century, the personal advertisements section in newspapers would sometimes be used to exchange messages encrypted using simple cipher schemes. Kahn (1967) describes instances of lovers engaging in secret communications enciphered using the Caesar cipher in *The Times*. Even as late as 1915, the Caesar cipher was in use: the Russian army employed it as a replacement for more complicated ciphers which had proved to be too difficult for their troops to master; German and Austrian cryptanalysts had little difficulty in decrypting their messages.

In April 2006, fugitive Mafia boss Bernardo Provenzano was captured in Sicily partly because of cryptanalysis of his messages written in a variation of the Caesar cipher. Provenzano's cipher used numbers, so that "A" would be written as "4", "B" as "5", and so on.

Breaking the cipher

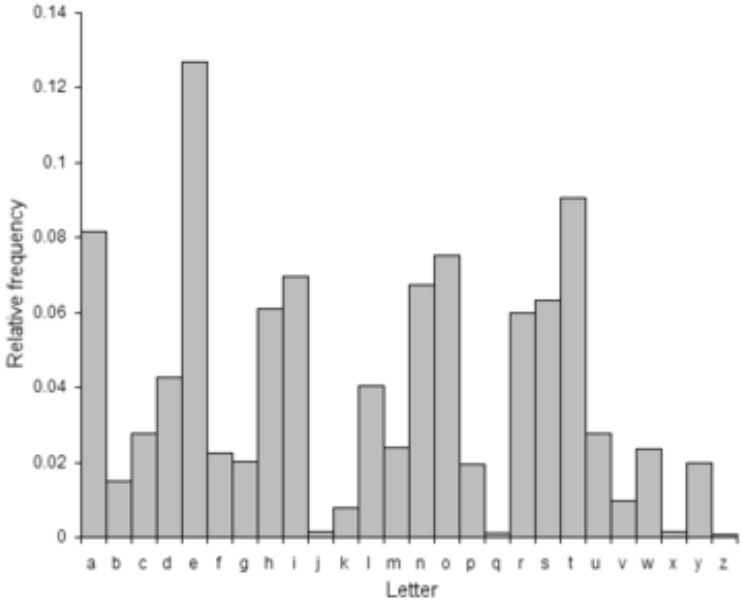
The Caesar cipher can be easily broken even in a ciphertext-only scenario using a frequency analysis.

Elements of probability provide tools for cryptanalysis, or breaking ciphers. Natural languages have characteristic letter frequencies, and these often show up directly in ciphertexts. Such characteristics can be exploited to break various types of substitution ciphers.

Frequency analysis is based on the fact that, in any given stretch of written language, certain letters and combinations of letters occur with varying frequencies. Moreover, there is a characteristic distribution of letters that is roughly the same for almost all samples of that language. For instance, given a section of English language, E tends to be very common, while X is very rare. Likewise, ST, NG, TH, and QU are common pairs of letters (termed *bigrams* or *digraphs*), while NZ and QJ are rare. The nonsense phrase "ETAOIN SHRDLU" represents the 12 most frequent letters in typical English language text.

In some ciphers, such properties of the natural language plaintext are preserved in the ciphertext, and these patterns have the potential to be exploited in a ciphertext-only attack

Relative frequencies of letters in the English language



Letter	Frequency				
a	8.167%	j	0.153%	t	9.056%
b	1.492%	k	0.772%	u	2.758%
c	2.782%	l	4.025%	v	0.978%
d	4.253%	m	2.406%	w	2.360%
e	12.702%	n	6.749%	x	0.150%
f	2.228%	o	7.507%	y	1.974%
g	2.015%	p	1.929%	z	0.074%
h	6.094%	q	0.095%		
i	6.966%	r	5.987%		
		s	6.327%		

In English, the space is slightly (107%) more frequent than the top letter, and the non-alphabetic characters (digits, punctuation, etc.) occupy the fourth position, between T and A.

Bigram

Bigrams or **digrams** are groups of two written letters, two syllables, or two words, and are very commonly used as the basis for simple statistical analysis of text. They are used in one of the most successful language models for speech recognition. They are a special case of N-gram.

The term is also used in cryptography, where *bigram frequency attacks* have sometimes been used to attempt to solve cryptograms.

Bigrams help provide the conditional probability of a word given the preceding word, when the relation of the conditional probability is applied:

$$P(W_n|W_{n-1}) = \frac{P(W_{n-1}, W_n)}{P(W_{n-1})}$$

That is, the probability $P()$ of a word W_n given the preceding word W_{n-1} is equal to the probability of their bigram, or the co-occurrence of the two words $P(W_{n-1},W_n)$, divided by the probability of the preceding word.

Bigram Frequency in the English language

The most common letter bigrams in the English language are listed below, with the expected number of occurrences per 2000 letters. In the analysis here, the bigrams are not permitted to span across consecutive words.

TH 50	AT 25	ST 20	HE 33	NT 24	AR 16
ER 40	EN 25	IO 18	IN 31	EA 22	AS 16
ON 39	ES 25	LE 18	ED 30	TI 22	DE 16
AN 38	OF 25	IS 17	ND 30	TO 22	RT 16
RE 36	OR 25	OU 17	HA 26	IT 20	VE 16

Two situations can be considered:

1. An attacker knows (or guesses) that some sort of simple substitution cipher has been used, but not specifically that it is a Caesar scheme;
2. An attacker knows that a Caesar cipher is in use, but does not know the shift value.

In the first case, the cipher can be broken using the same techniques as for a general simple substitution cipher, such as frequency analysis or pattern words. While solving, it is likely that an attacker will quickly notice the regularity in the solution and deduce that a Caesar cipher is the specific algorithm employed.

In the second instance, breaking the scheme is even more straightforward. Since there are only a limited number of possible shifts (26 in English), they can each be tested in turn in a brute force attack. One way to do this is to write out a snippet of the ciphertext in a table of all possible shifts — a technique sometimes known as "completing the plain component". The example given is for the ciphertext "EXXEGOEXSRGI"; the plaintext is instantly recognisable by eye at a shift of four. Another way of viewing this method is that, under each letter of the ciphertext, the entire alphabet is written out in reverse starting at that letter. This attack can be accelerated using a set of strips prepared with the alphabet written down them in reverse order. The strips are then aligned to form the ciphertext along one row, and the plaintext should appear in one of the other rows.

Another brute force approach is to match up the frequency distribution of the letters. By graphing the frequencies of letters in the ciphertext, and by knowing the expected distribution of those letters in the original language of the plaintext, a human can easily spot the value of the shift by looking at the displacement of particular features of the graph. This is known as frequency analysis. For example in the English language the plaintext frequencies of the letters E, T, (usually most frequent), and Q, Z (typically least frequent) are particularly distinctive. Computers can also do this by measuring how well the actual frequency distribution matches up with the expected distribution; for example, the chi-square statistic can be used.

For natural language plaintext, there will, in all likelihood, be only one plausible decryption, although for extremely short plaintexts, multiple candidates are possible. For example, the ciphertext MPQY could, plausibly, decrypt to either "aden" or "know" (assuming the plaintext is in English); similarly, "ALIIP" to "dolls" or "wheel"; and "AFCCP" to "jolly" or "cheer" (see also unicity distance).

Multiple encryptions and decryptions provide no additional security. This is because two encryptions of, say, shift *A* and shift *B*, will be equivalent to an encryption with shift *A* + *B*. In mathematical terms, the encryption under various keys forms a group.

2.4. The Playfair cipher

The Playfair cipher or Playfair square is a manual symmetric encryption technique and was the first literal digraph substitution cipher. The scheme was invented in 1854 by Charles Wheatstone, but bears the name of Lord Playfair who promoted the use of the cipher.The Playfair is thus significantly harder to break since the frequency analysis used for simple substitution ciphers does not work with it. Frequency analysis can still be undertaken, but on the 600 possible digraphs rather than the 26 possible monographs. The frequency analysis of digraphs is possible, but considerably more difficult – and it generally requires a much larger ciphertext in order to be useful.

To generate the key table, one would first fill in the spaces in the table with the letters of the keyword (dropping any duplicate letters), then fill the remaining spaces with the rest of the letters of the alphabet in order (usually omitting "Q" to reduce the alphabet to fit, other versions put both "I" and "J" in the same space). The key can be written in the top rows of the table, from left to right, or in some other pattern, such as a spiral beginning in the upper-left-hand corner and ending in the center. The keyword together with the conventions for filling in the 5 by 5 table constitute the cipher key.

To encrypt a message, one would break the message into digraphs (groups of 2 letters) such that, for example, "HelloWorld" becomes "HE LL OW OR LD", and map them out on the key table. The two letters of the digraph are considered as the opposite corners of a rectangle in the key table. Note the relative position of the corners of this rectangle. Then apply the following 4 rules, in order, to each pair of letters in the plaintext:

- If both letters are the same (or only one letter is left), add an "X" after the first letter. Encrypt the new pair and continue.

- If the letters appear on the same row of your table, replace them with the letters to their immediate right respectively (wrapping around to the left side of the row if a letter in the original pair was on the right side of the row).
- If the letters appear on the same column of your table, replace them with the letters immediately below respectively (wrapping around to the top side of the column if a letter in the original pair was on the bottom side of the column).
- If the letters are not on the same row or column, replace them with the letters on the same row respectively but at the other pair of corners of the rectangle defined by the original pair. The order is important – the first letter of the encrypted pair is the one that lies on the same **row** as the first letter of the plaintext pair.

To decrypt, use the INVERSE (opposite) of the first 3 rules, and the 4th as is (dropping any extra "X"s that don't make sense in the final message when you finish).

Example

Using "playfair example" as the key, the table becomes:

Encrypting the message "Hide the gold in the tree stump":

HI DE TH EG OL DI NT HE TR EX ES TU MP

P L A Y F
I R E X M
B C D G H
K N O Q S
T U V W Z

The pair HI forms a rectangle, replace it with BM

1. The pair DE is in a column, replace it with OD
2. The pair TH forms a rectangle, replace it with ZB
3. The pair EG forms a rectangle, replace it with XD
4. The pair OL forms a rectangle, replace it with NA
5. The pair DI forms a rectangle, replace it with BE
6. The pair NT forms a rectangle, replace it with KU
7. The pair HE forms a rectangle, replace it with DM
8. The pair TR forms a rectangle, replace it with UI
9. The pair EX (X inserted to split EE) is in a row, replace it with XM
10. The pair ES forms a rectangle, replace it with MO
11. The pair TU is in a row, replace it with UV
12. The pair MP forms a rectangle, replace it with IF

BM OD ZB XD NA BE KU DM UI XM MO UV IF

Thus the message "Hide the gold in the tree stump" becomes
"BMODZBXDNABEKUDMUIXMMOUVIF".

Clarification with pictures

Assume one wants to encrypt the digraph OR. There are three general cases:

1)	2)	3)
* * * * *	* * O * *	Z * * O *
* O Y R Z	* * B * *	* * * * *
* * * * *	* * * * *	* * * * *
* * * * *	* * R * *	R * * X *
* * * * *	* * Y * *	* * * * *
Hence, OR -> YZ	Hence, OR -> BY	Hence, OR -> ZX

Playfair cryptanalysis

Like most pre-modern era ciphers, the Playfair cipher can be easily cracked if there is enough text. Obtaining the key is relatively straightforward if both plaintext and ciphertext are known. When only the ciphertext is known, brute force cryptanalysis of the cipher involves searching through the key space for matches between the frequency of occurrence of digrams (pairs of letters) and the known frequency of occurrence of digrams in the assumed language of the original message.

- Cryptanalysis of Playfair is similar to that of four-square and two-square ciphers, though the relative simplicity of the Playfair system makes identifying candidate plaintext strings easier. Most notably, a Playfair digraph and its reverse (e.g. AB and BA) will decrypt to the same letter pattern in the plaintext (e.g. RE and ER). In English, there are many words which contain these reversed digraphs such as REceivER and DEpartED. Identifying nearby reversed digraphs in the ciphertext and matching the pattern to a list of known plaintext words containing the pattern is an easy way to generate possible plaintext strings with which to Topics in cryptography

Notes

No duplicate letters are allowed, and one letter is omitted (Q) or combined (I/J), so the calculation is $600 = 25 \times 24$.

For Web for interactive cooding:

see: http://www.simonsingh.net/The_Black_Chamber/playfaircipher.htm

2.5. ADFGVX Cipher (1918 WW I)

ADFGVX Cipher was used by the German Army during World War I. This cipher was restricted to German High Command communications between and among the headquarters of divisions and army corps.

One weakness of a Caeser Cypher cipher is that is does not hide the frequency of the letters that appear in the natural language. It just changes the symbol used for the letter. ADFGVX Cipher overcomes this by replacing a single letter with a pair of letters then scrambling the pairs in a manner similar to the Spartan Syctale.

Step 1 of Encoding
To encrypt a message with the ADFGVX cipher there are two steps. First uses the ADFGVX grid to replace letters of plain text with a pair of letters that correspond to the row and column of the letter in a grid. Spaces are removed since the grid does not include spaces.

Examples:

- K=>“FD”
- A=>“DV”
- Kaiser Wilhelm=>“FDDVVGVDXDVVGGVGDADXXDDAGX”

The particular letters, ADFGVX, were chosen for their distinctive Morse encoding to prevent errors in transmission.

	A	**D**	**F**	**G**	**V**	**X**
A	8	P	3	D	1	N
D	L	T	4	O	A	H
F	7	K	B	C	5	Z
G	J	U	6	W	G	M
V	X	S	V	I	R	2
X	9	E	Y	0	F	Q

Step 2 of Encoding

The second step is similar to the Spartan Scytale except there is a key word that scrambles the order further. Suppose the key word is “RHIEN”. The encoded word is placed in a table with as many columns as the key word using row-major ordering. The columns are then rearranged by the alphabetical order of the key word’s letters and the final encoding is then read from the table in a column-major order. Finally, spaces are added to aid readability for the person transmitting or receiving the encrypted text.

R	**H**	**I**	**E**	**N**
5	**2**	**3**	**1**	**4**
F	**D**	**D**	**V**	**V**
G	**V**	**D**	**X**	**D**
V	**V**	**G**	**G**	**V**
G	**D**	**A**	**D**	**X**
X	**D**	**D**	**A**	**G**
X				

Final Encoding of “Kaiser Wilhelm”: VXGDADVVDDDDGADV DVXG FGVG XX

2.6. Vigenere Square (1585)

The Vigenère cipher uses a Caesar cipher with a different shift at each position in the text; the value of the shift is defined using a repeating keyword. If a single-use keyword is as long as the message and chosen randomly then this is a one-time pad cipher, unbreakable if the users maintain the keyword's secrecy. Keywords shorter than the message (e.g., "Complete Victory" used by the Confederacy during the American Civil War), introduce a cyclic pattern that might be detected with a statistically advanced version of frequency analysis.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Ciphers base on a Vigenere Square are similar to a Caersar Cipher with variable shifts. The shifts are based on a key that is known only to the encoder and decoder. It works this way. Suppose the key is the word RAIN and you wish to encode THISISGREAT

Here is how it is done:
Replace “T” with “T-R” entry “J”.

PT	T	H	I	S	I	S	G	R	E	A	T
Key	***R***	***A***	***I***	***N***	***R***	***A***	***I***	***N***	***R***	***A***	***I***
CT	J	H	Q	E	Z	.	.	.	.	.	.

Replace “H” with “H-A” entry “H”.
Replace “I” with “I-I” entry “Q”.
Replace “S” with “S-N” entry “E”.
Replace “I” with “I-R” entry “Z”
Etc

2.7. Nihilist cipher

In the history of cryptography, the **Nihilist cipher** is a manually operated symmetric encryption cipher originally used by Russian Nihilists in the 1880s to organize terrorism against the czarist regime. Both the plaintext and a keyword is converted to a series of two digit numbers. These numbers are then added together in the normal way to get the ciphertext, with the key numbers repeated as required.

Example

Consider the Polybius square created using the keyword ZEBRAS:

	1	**2**	**3**	**4**	**5**
1	Z	E	B	R	A
2	S	C	D	F	G
3	H	I	K	L	M
4	N	O	P	Q	T
5	U	V	W	X	Y

with a plaintext of "DYNAMITE WINTER PALACE" and a key of RUSSIAN. This expands to:

```
PT:     23  55  41  15  35  32  45  12  53  32  41  45  12  14  43  15  34  15  22  12
KEY:    14  51  21  21  32  15  41  14  51  21  21  32  15  41  14  51  21  21  32  15
CT:     37  106 62  36  67  47  86  26  104 53  62  77  27  55  57  66  55  36  54  27
```

2.8. Transposition cipher

In cryptography, a **transposition cipher** is a method of encryption by which the positions held by units of plaintext (which are commonly characters or groups of characters) are shifted according to a regular system, so that the ciphertext constitutes a permutation of the plaintext. That is, the order of the units is changed. Mathematically a bijective function is used on the characters' positions to encrypt and an inverse function to decrypt.

Following are some implementations.

Rail Fence cipher

The Rail Fence cipher is a form of transposition cipher that gets its name from the way in which it is encoded. In the rail fence cipher, the plaintext is written downwards on successive "rails" of an imaginary fence, then moving up when we get to the bottom. The message is then read off in rows. For example, using three "rails" and a message of 'WE ARE DISCOVERED. FLEE AT ONCE', the cipherer writes out:

Size of Shift : 3

W R I O R . E T C

E E S V E F E O E

A D C E D L A N

Then reads off:

WECRL TEERD SOEEF EAOCA IVDEN

(The cipherer has broken this ciphertext up into blocks of five to help avoid errors.)

Online version: http://www.simonsingh.net/The_Black_Chamber/railfence.html

Route cipher

In a route cipher, the plaintext is first written out in a grid of given dimensions, then read off in a pattern given in the key. For example, using the same plaintext that we used for rail fence:

W R I O R F E O E
E E S V E L A N J
A D C E D E T C X

The key might specify "spiral inwards, clockwise, starting from the top right". That would give a cipher text of:

EJXCTEDECDAEWRIORFEONALEVSE

Route ciphers have many more keys than a rail fence. In fact, for messages of reasonable length, the number of possible keys is potentially too great to be enumerated even by modern machinery. However, not all keys are equally good. Badly chosen routes will leave excessive chunks of plaintext, or text simply reversed, and this will give cryptanalysts a clue as to the routes.

An interesting variation of the route cipher was the Union Route Cipher, used by Union forces during the American Civil War. This worked much like an ordinary route cipher, but transposed whole words instead of individual letters. Because this would leave certain highly sensitive words exposed, such words would first be concealed by code. The cipher clerk may also add entire null words, which were often chosen to make the ciphertext humorous.

Columnar transposition

In a columnar transposition, the message is written out in rows of a fixed length, and then read out again column by column, and the columns are chosen in some scrambled order. Both the width of the rows and the permutation of the columns are usually defined by a keyword. For example, the word ZEBRAS is of length 6 (so the rows are of length 6), and the permutation is defined by the alphabetical order of the letters in the keyword. In this case, the order would be "6 3 2 4 1 5".

In a regular columnar transposition cipher, any spare spaces are filled with nulls; in an irregular columnar transposition cipher, the spaces are left blank. Finally, the message is read off in columns, in the order specified by the keyword. For example, suppose we use the keyword ZEBRAS and the message WE ARE DISCOVERED FLEE AT ONCE. In a regular columnar transposition, we write this into the grid as:

6	3	2	4	1	5
W	E	A	R	E	D
I	S	C	O	V	E
R	E	D	F	L	E
E	A	T	O	N	C
E	Q	K	J	E	U

Providing five nulls (QKJEU) at the end. The ciphertext is then read off as:

EVLNE ACDTK ESEAQ ROFOJ DEECU WIREE

In the irregular case, the columns are not completed by nulls:

6	3	2	4	1	5
W	E	A	R	E	D
I	S	C	O	V	E
R	E	D	F	L	E
E	A	T	O	N	C
E					

This results in the following ciphertext:

EVLNA CDTES EAROF ODEEC WIREE

To decipher it, the recipient has to work out the column lengths by dividing the message length by the key length. Then he can write the message out in columns again, then re-order the columns by reforming the key word.

Columnar transposition continued to be used for serious purposes as a component of more complex ciphers at least into the 1950's.

Double transposition

A single columnar transposition could be attacked by guessing possible column lengths, writing the message out in its columns (but in the wrong order, as the key is not yet known), and then looking for possible anagrams. Thus to make it stronger, a double transposition was often used. This is simply a columnar transposition applied twice. The same key can be used for both transpositions, or two different keys can be used.

As an example, we can take the result of the irregular columnar transposition in the previous section, and perform a second encryption with a different keyword, STRIPE, which gives the permutation "564231":

5	6	4	2	3	1
E	V	L	N	A	C
D	T	E	S	E	A
R	O	F	O	D	E
E	C	W	I	R	E
E					

As before, this is read off columnwise to give the ciphertext:

CAEEN SOIAE DRLEF WEDRE EVTOC

During World War I, the German military used a double columnar transposition cipher. The system was regularly solved by the French, naming it Übchi, who were typically able to find the key in a matter of days after a new one had been introduced. However, the French success became widely-known and, after a publication in *Le Matin*, the Germans changed to a new system on 18 November 1914.

During World War II, the double transposition cipher was used by Dutch Resistance groups, the French Maquis and the British Special Operations Executive (SOE), which was in charge of managing underground activities in Europe. It was also used by agents of the American Office of Strategic Services and as an emergency cipher for the German Army and Navy.

Until the invention of the VIC cipher, double transposition was generally regarded as the most complicated cipher that an agent could operate reliably under difficult field conditions.

Myszkowski transposition

A variant form of columnar transposition, proposed by Émile Victor Théodore Myszkowski in 1902, requires a keyword with recurrent letters. In usual practice, subsequent occurrences of a keyword letter are treated as if the next letter in alphabetical order, *e.g.,* the keyword TOMATO yields a numeric keystring of "532164."
In Myszkowski transposition, recurrent keyword letters are numbered identically, TOMATO yielding a keystring of "432143."

4	3	2	1	4	3
W	E	A	R	E	D
I	S	C	O	V	E
R	E	D	F	L	E
E	A	T	O	N	C

E

Plaintext columns with unique numbers are transcribed downward; those with recurring numbers are transcribed left to right:

ROFOA CDTED SEEEA CWEIV RLENE

Bifid Cipher

The Bifid cipher is a type of matrix, or columnar transposition, cipher.
Start by creating a 5 by 5 matrix of letters, with the rows and columns labeled 1 to 5.

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I	J
3	K	L	M	N	O
4	P	Q	R	S	T
5	U	V	W	X	YZ

To start, find the value of each letter by reading the row and the column values. The two numbers are then written vertically on a piece of paper below the plain letter. All the plain letters within the secret message are written next to one another as seen below:

Plain Message: S E N D R E I N F O R C E M E N T

Row Value: 4 1 3 1 4 1 4 3 2 3 4 1 1 3 1 3 4

Column Value: 4 5 4 4 3 5 4 4 1 5 3 3 5 3 5 4 5

Notice how the letter "S" has the value of 44. "E" is 15 since it is found in row 1, column 5. Y and Z share the position of (5,5) in the matrix above. After the message has been written out, with row and column values written as shown above, you rewrite the message from left to right, combing numbers into groups of 2.

41 31 41 43 23 41 13 13 44 54 43 54 41 53 35 35 45

The last step is to take each group of numbers, such as 41 and 31 in the beginning of the line above, and find the corresponding cipher values in the same matrix above. 41 is row 4, column 1, the letter "P."

41 31 41 43 23 41 13 13 44 54 43 54 41 53 35 35 45

P K P R H P C C S X R X P W O O T

Detection and cryptanalysis

Since transposition does not affect the frequency of individual symbols, simple transposition can be easily detected by the cryptanalyst by doing a frequency count. If the ciphertext exhibits a frequency distribution very similar to plaintext, it is most likely a transposition. This can then often be attacked by anagramming - sliding pieces of ciphertext around, then looking for sections that look like anagrams of English words, and solving the anagrams. Once such anagrams have been found, they reveal information about the transposition pattern, and can consequently be extended.

Simpler transpositions also often suffer from the property that keys very close to the correct key will reveal long sections of legible plaintext interspersed by gibberish. Consequently such ciphers may be vulnerable to optimum seeking algorithms such as genetic algorithms.

Combinations

Transposition is often combined with other techniques. For example, a simple substitution cipher combined with a columnar transposition avoids the weakness of both. Replacing high frequency ciphertext symbols with high frequency plaintext letters does not reveal chunks of plaintext because of the transposition. Anagramming the transposition does not work because of the substitution. The technique is particularly powerful if combined with fractionation (see below). A disadvantage is that such ciphers are considerably more laborious and error prone than simpler ciphers.

Fractionation

Transposition is particularly effective when employed with fractionation - that is, a preliminary stage that divides each plaintext symbol into several ciphertext symbols. For example, the plaintext alphabet could be written out in a grid, then every letter in the message replaced by its co-ordinates (see Polybius square). Another method of fractionation is to simply convert the message to Morse code, with a symbol for spaces as well as dots and dashes.

When such a fractionated message is transposed, the components of individual letters become widely separated in the message, thus achieving Claude E. Shannon's diffusion. Examples of ciphers that combine fractionation and transposition include the bifid cipher, the trifid cipher, the ADFGVX cipher and the VIC cipher.

Another choice would be to replace each letter with its binary representation, transpose that, and then convert the new binary string into the corresponding ASCII characters. Looping the scrambling process on the binary string multiple times before changing it into ASCII characters would likely make it harder to break. Many modern block ciphers use more complex forms of transposition related to this simple idea.